R E P O R T R E S U M E S

ED 010 361 48

LANGUAGES OF THE WORLD--SINO-TIBETAN FASCICLE FOUR.

BY- VOEGELIN, C. F. VOEGELIN, FLORENCE M.
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER NDEA-VI-63-14 PUB DATE MAY 65
CONTRACT OEC-SAE-9488
EDRS PRICE MF-$0.18 HC-$2.52 63P. ANTHROPOLOGICAL LINGUISTICS, 7(5)/1-57, MAY 1965

DESCRIPTORS- *LANGUAGES, *SINO TIBETAN LANGUAGES, *CHINESE, ARCHIVES OF LANGUAGES OF THE WORLD, BLOOMINGTON, INDIANA

THE LANGUAGES OF THE TIBETAN FAMILY PROPER ARE LISTED AND DESCRIBED. (THIS REPORT IS PART OF A SERIES, ED 010 350 TO ED 010 367.) (JK)

ED010361

63-14
SAE-9488 OF

PART I

Anthropological Linguistics

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Volume 7 Number 5

May 1965

title

LANGUAGES OF THE WORLD:

SINO-TIBETAN FASCICLE FOUR

A Publication of the
ARCHIVES OF LANGUAGES OF THE WORLD
Anthropology Department
Indiana University

ANTHROPOLOGICAL LINGUISTICS is designed primarily, but not exclusively, for the immediate publication of data-oriented papers for which attestation is available in the form of tape recordings on deposit in the Archives of Languages of the World. This does not imply that contributors will be restricted to scholars working in the Archives at Indiana University; in fact, one motivation for the publication of ANTHROPOLOGICAL LINGUISTICS is to increase the usefulness of the Archives to scholars elsewhere by making publishable data and conclusions and their tape recorded attestation more widely available. (Recorded attestation of papers from scholars elsewhere will be copied by the Archives and the original recordings returned to the collector; others may then work with the tape copy either in the Archives or elsewhere by having a copy sent to them.) In addition to heavily exemplified papers in the form of preliminary or final statements on restricted problems in phonemics, morphophonemics, morphemics, syntax and comparative grammar, ANTHROPOLOGICAL LINGUISTICS will include lexical lists and analyzed texts (especially in the otherwise hard-to-publish range of 20 to 100 pages) and theoretical and methodological papers (especially in the form of papers from symposia).

Each volume of ANTHROPOLOGICAL LINGUISTICS consists of nine numbers to be issued during the months of January, February, March, April, May, June, October, November and December. Subscriptions ($3.50 a year) and papers for publication should be sent to the editor, Dr. Florence M. Voegelin, Anthropology Department, Indiana University, Bloomington, Indiana.

Entered as second class matter at the postoffice
at Bloomington, Indiana

LANGUAGES OF THE WORLD:
SINO-TIBETAN FASCICLE FOUR

C. F. and F. M. Voegelin

Indiana University

4.0. Scope of the Tibetan family

4.1. Sample of Tibetan sentences

4.2. West Tibetan

4.3. Lhasa Tibetan

4.4. Non-Lhasa Central Tibetan

5.0. Scope of the Gyarung-Mishmi family

5.1. Western complex languages

5.2. 'Pronominalized' languages

5.3. 'Non-pronominalized' languages

For authorship and sponsorship, see Languages of the World; Sino-Tibetan Fascicle One (0.1). The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education and Welfare.

4.0. Classical Tibetan provides a standard or point of reference for many students of the Tibetan languages as they are spoken today, much as Classical Arabic—the Arabic in which the Koran is written—serves as a point of departure for some students of modern Arabic dialects and for some comparative studies of other Semitic languages. The lexical resources of Classical Tibet are brought together in a Tibetan Word Book, to which Gould and Richardson (1943) contribute historical perspective. There is a tradition of an original Tibetan homeland, called Amdo. There is evidence of extensive Tibetan contact with India and China. In the 7th century, Buddhism was made the state religion, which stimulated the adaptation of a writing system and translation of Buddhistic literature from Sanskrit to Tibetan. Morphemes in Sanskrit are often polysyllabic but, it is said, Tibetan morphemes are virtually all monosyllabic. "In Tibetan almost every syllable has a clear and definite meaning." The Tibetan Word Book places two thousand of these 'syllables'—i.e. morphemes—in Tibetan alphabetic order, and cites the morphemes in Tibetan script, together with a transliteration of the script, and a phonetic transcription of the Modern Lhasa Tibetan equivalent of each morpheme.

Aside from the single usage for Classical Tibetan, the literature on this subject reflects two meanings in the use of 'Tibetan' for languages which are spoken today; this dual usage can be confusing if it is not understood at the outset. On the one hand, 'Tibetan' is used as a cover-term for the Tibetan languages proper (see below), which are the descendants of Classical Tibetan. Some of the literature, however, divides the Sino-Tibetan phylum into an East Asia part and an extended South Asia or Tibeto-Burman part. When Chinese scholars divide the Sino-Tibetan phylum, they tend to

make four divisions: Chinese (our Han Chinese family, 1, above), Kam-Thai (2, above), Miao-Yao (3, above), and all the rest, called Tibeto-Burman. In the dual usage found in the literature, the term 'Tibetan' is used not only for languages of the Tibetan family, but also for any non-Burman Tibeto-Burman language—that is to say, some of the literature uses 'Tibetan' as a classificatory term to include Tibetan proper, the so-called Himalayan languages (e.g. Gyarung), and the so-called North Assam languages (e.g. Mishmi); so also, as a parallel classificatory term, 'Burman' is used to include Burmese-Lolo, Bodo, $Naga_1$, Kachin, $Naga_2$, and Chin languages. We do not employ 'Burman' as a classificatory term. We avoid the danger of confusion or ambiguity in the use of the term 'Tibetan' by using it exclusively in reference to languages of the Tibetan family. We follow the usage of Chinese scholars in the first three divisions or language families, of the Sino-Tibetan phylum (Chinese, Kam-Thai, Miao-Yao), but instead of grouping all the rest of Sino-Tibetan that is found west of these three language families into Tibeto-Burman, we distinguish coordinately the Tibetan family of languages, from the Burmese-Lolo family, from the Karen family, from the Bodo-$Naga_1$-Kachin family, from the $Naga_2$-Chin family, and from the Gyarung-Mishmi family.

The Tibetan family proper includes only two or three separate languages but many dialects. In his comprehensive survey of this family, Grierson (1909) lists language and dialect names, without distinguishing between the two, in terms of the relative geographic location of the speakers, thereby, obtaining a Western, Central, and Eastern group of Tibetan speakers. The self-designation for 'Tibetan' or 'Tibetan of' a given place is Bhōṭia.

Under Western Tibetan, there are listed:

Baltī

Purik

Ladakhī

Lahul.

Under Central Tibetan, there are listed:

Central dialect (see below)

Spiti

Nyamkat

Jad

Garhwal

Kāgate

Sharpa

Dänjong-Kä (Bhōtia of Sikkim)

Lhoke (Bhōtia of Bhutan).

Under Eastern Tibetan, there is listed:

Kham,

which is said to share the same general structure reflected by Western Tibetan languages. Recent scholars do not distinguish any 'Eastern' branch which was recognized earlier as having no linguistic basis for separate classification; hence Kham is included below as one dialect of a language whose other dialects are spoken much to the west of Kham.

The various statements of mutual intelligibility between the many names given suggest that there are few language barriers in the Tibetan family, but much dialect differentiation. Thus, the so-called 'Eastern dialects' and all the 'Western dialects' may constitute one language with no tone distinctions, but with prefixes preserved; apparently all dialects in the West Tibetan language share some such features of sound (no tone) and word

structure (prefixes). This West Tibetan language, spoken in many dialects, must then be distinguished from at least one other Tibetan language, tentatively called the Central Tibetan language, including Lhasa and non-Lhasa dialects. There are groups of dialects within the Central Tibetan language, and there is also a criss-crossing classification of Lhasa vs. non-Lhasa dialects. Some 'Central' appears as though it referred to (1) all those dialects of the Central Tibetan language that remain after subtracting (2) the closely similar dialects of Spiti, Nyamkat, Jad, etc., and (3) the closely similar dialects of Kagate, Sharpa, etc. It might turn out of course that this three-way division of dialects represents three separate languages which together with the West Tibetan language would make a total of four languages for the Tibetan family, rather than the two, the Central Tibetan language and the West Tibetan language, which data now available serves to support.

A very clear generalization by Grierson (1909) states that dialects of the Central Tibetan language have developed tones but have lost prefixes, and have merged original voiced consonant initials with voiceless consonants, either aspirated or unaspirated. Generalizations like this can be checked against the information on specific dialects from the Central Tibetan language, as well as from the West Tibetan language (4.2, below).

A more recent generalization by Roy Andrew Miller (1955) distinguishes typologically between (1) 'simple' or 'innovating' Tibetan dialects (meaning thereby those dialects which are now spoken without consonant clusters either as initials or as finals, but with morphemes distinguished by tones) and (2) 'complex' or 'archaic' Tibetan dialects (meaning thereby those dialects which are now spoken with consonant clusters as initials and also as finals, but with morphemes not distinguished by differences in tone). These typological

criteria for dialects of type (1) and type (2) correlate with the minimum of two separate languages in the Tibetan family, mentioned above. Dialects of the West Tibetan language reflect type (2) — 'complex' or 'archaic'; dialects of the Central Tibetan language reflect type (1) — 'simple' and 'innovating'.

4.1. The following sentences are examples of the Central Tibetan dialect proper, one of many Tibetan dialects (4.2 to 4.4, below). In a general statement comparing relative order of Tibetan with Chinese and Tai, Grierson (1909), points out that the latter favor a S-V-O order (subject-verb-object), while Tibetan favors a S-O-V order (subject-object-verb). In Chinese and Tai, as well as in Central Tibetan, the object as topic may precede the verb or subject-verb (O-V or O-(S-V) order). These relative orders are only favorite orders of clauses, or more generally, of topic and comment, which may include more than one phrase, and the phrases may often be reordered in non-contrastive syntax, as has been shown in our sample of Chinese and in our sample of Thai sentences where our focus of attention was on the order of morphemes within the phrase. In the sample of Tibetan sentences we cite instances of O-S-V order (object-subject-verb) but focus attention on the phrases which include morphemes which are not reorderable in a non-contrastive way, as the phrases themselves are. The relative orders of morphemes within a phrase are not merely favorite orders, as some of the relative orders among phrases may be, but invariable orders for marking relationships. Thus, within a phrase, modified precedes modifier (noun-adjective) both in Tibetan and Thai, while the Chinese fixed order is modifier before modified (adjective-noun).

Under the Tibetan sentences, which are first freely translated, there is a literal translation or gloss of each successive morpheme. Square brackets enclose glosses of the morphemes above them which are included in phrases. Every major morpheme (capital M) can function alone as a 'word' or phrase nucleus. In our sample there are 45 minor morphemes (abbreviated as lower-case m, in the discussion after some sentences), none of which ever function as a phrase nucleus; some of these can be classified as affixes, as dependent parts of a word. (The formal criteria for a word in Tibetan, as given by Spriggs (1955), cannot always be applied to our sentences for lack of phonetic information.) In our sample, there are 25 dual function morphemes ($\frac{m}{M} \sim \frac{M}{m}$), beside the 45 minor morphemes that never function as phrase nucleus (lower-case m's), and beside uncounted but numerous major morphemes (M). A dual function ($\frac{m}{M} \sim \frac{M}{m}$) morpheme may function as phrase nucleus in one phrase ($\frac{M}{m}$); the same morpheme, in another phrase, may function in a dependency relationship ($\frac{m}{M}$) to a major morpheme in the same phrase. Where a given morpheme is hard to gloss clearly, the minor morpheme symbol (m or $\frac{m}{M}$) is given in the interlinear bracket phrasing, and the literal translation is given at length right after the sentence. When a 'word' is written as a sequence of major morpheme and minor morpheme (M-m), with the minor morpheme functioning as a suffix, the 'word' is hyphenated, and each morpheme in the 'word' (M and m) is glossed separately.

I am not able to say whether this is good or not.

(1) di yak po yī̱ m-ē̱ ŋ-ē̄

[this good adj. m not-m] [I-m]

lap t'u pa min-du

[say able pf. not-present]

The first phrase, glossed [this good adj.], shows the modified-modifier order. The morpheme for this (M) is modified by the morpheme for good (M) which is followed by a morpheme glossed adjective or adjective formative (m); the next m cited in this phrase means non-present, and appears again suffixed to the morpheme for not (in an alternate shape), so that the phrase as a whole means whether this is good or not good. The second phrase includes two morphemes, glossed I-m to show that a suffix follows the first person actor marker; the suffix indicates that the actor is actor of a transitive verb. The third phrase begins with a major morpheme for say followed by verb complements for able and perfective, followed by the morpheme for not followed by complement marking present, so that not able to say is expressed as a present perfective.

There is one palace inside the town.

(2) toŋ pēi k'i-la p'otaŋ čǐ du

[town of inside-in][palace one] [present]

The second phrase, glossed [palace one], shows again the modified-modifier order. The third phrase includes the morpheme for present non-first person, functioning here as the phrase nucleus or verb, is.

The soldier shot at the enemy.

(3) ma mi ḍa la mendā č'ap pa re

[(war man)] [enemy goal] [(rifle throw) perfective non-1st-person]

The first phrase gives the subject of the sentence, by a compound indicated in notation as (M M), or by putting in parentheses the glosses for the two major morphemes of the compound. The second phrase specifies the object of the sentence redundantly—by its favorite order position (after subject and before verb,

S-O-V), and by a minor morpheme which functions as a goal marker after the major morpheme for enemy within the phrase. In the third or verb phrase, the compound verb (rifle throw) for shoot, is followed by two verbal complements--that for perfective is always minor morpheme (m), while that for non-1st-person is a dual function morpheme, minor when a verb complement ($\underline{\underline{m}}$ (M)), but major ($\underline{\underline{M}}$ (m)) when phrase nucleus.

I very much want to go.

(4) ŋ-a ḍo-ndö šeṭa yö

[I-m] [go-want exceedingly $\frac{m}{M}$]

The first phrase includes two morphemes, glossed I-m to show that the morpheme for I is followed by a suffix (-m) which specifies that the actor is actor of an intransitive verb. The second phrase begins with verb nucleus morpheme for go before suffix for want. The next morpheme, marking very much, exceedingly, is modifier to the preceding go-want; more particularly, it is the morpheme for want that is modified. The last morpheme in this phrase is a minor morpheme here ($\frac{m}{M}$): a verbal complement which specifies redundantly the person of the actor (first person), as well as the present tense of the verb phrase.

I was working.

(5) ŋ-ā leka č'ē pa yö

[I-m] [(work do) perfective 1st-person-present]

In the second phrase, two verbal complements follow the phrase nucleus, a compound (hence in parentheses) which functions as a verb. Other compounds of noun-verb sequences that function as verbs are (rifle throw) for shoot, as in sentence (3), above; (water pour) for to water (č'u lū);(help do) for assist (rō č'ē). Greater variety appears in compounds that function as nouns,

as in sentence (3): (war man) for soldier, in which the first noun is in a relative-possessive if not exactly modifier relationship to the head noun; so also (king son) for prince (g'e se); (man lord) for lord of men (mi waŋ); road pass for passport (lam jik). Redundant compounds are those in which more or less synonymous nouns mark the common meaning, as (swamp mud) for mire (dam dzap); (well-being happy) for peace; and (color color) for color (c'ö̆ do). In other coordinate compounds, the polar extremes of the referent are specified by the members of the compound; and the referent of the compound can be predicted from the members more readily than is the case in such Chinese compounds as (east-west) for something; thus, the Central Tibetan compound for size is (big small, č'e č'uŋ); distance is (long short, riŋ t'uŋ); dampness is (dry wet, kam lo̱). There are other coordinate relationships, as in the compounds for flock (goat sheep, ra lu), and for transportation (horse mule, ta ṭa). The classifiers, which precede the appropriate nouns in a dozen or more domains, as /ku/ before nouns in the domain of body-parts, appear to be another but specialized class of compound members.

As for this work, I am not able to do it.

(6)	leka	di	ŋ-ɛ̄	č'e	t'u	pa	min-du
	[work	this]	[I-m	do	able	perfective	not-present]

The first phrase exemplifies again the modified-modifier order, and also shows the use of topic as goal or object of the sentence. The second phrase, the comment, begins with an embedded actor I-m, in which the morpheme for first person actor is followed by -m, a suffix indicating that this actor is the actor of a verb functioning as a transitive verb (but the object appears in the prior topic), followed by the verb for do, followed by verbal complements, as in sentence (1), above.

As for his very small horse, I killed it.

(7) k'-ö̆i ta č'uŋ č'uŋ ŋ-ē se pa yī̆
[he-of horse (small small)] [I-m] [kill perfective ($\frac{m}{M}$)]

In the first phrase, his is marked by morpheme for third person followed by genitive suffix, glossed of, before the modified morpheme horse, followed by reduplicated modifier (small small); here again the topic functions as goal or object of the sentence. The second phrase begins with embedded subject, I, with suffix indicating that 'I' is actor of a transitive verb, followed by the phrase nucleus, kill, before verb complements, the last of which redundantly marks 1st person actor, as well as non-present tense (past in this sentence, but future in other sentences). The reduplication of the modifier in the first phrase serves for emphasis. And so generally: (round round) for very round, /gō gō/; (flat flat) for very flat, /lep lep/.

That he was writing it, I saw for myself.

(8) k'-oŋ gī ṭī p-ā ŋ-a raŋ raŋ gī t'oŋ č'uŋ
[he-m m write -ing-goal] [I-m (self self) m see pf.]

In the second phrase, the phrase nucleus for see is followed by a verb complement for a perfective which is always past tense. The embedded actor precedes the morpheme for see; it is a person marker for I followed by suffix for actor of transitive verb and by reduplicated reflexive (self self) as modifier of the preceding person marker. The fact that the actor is the actor of a transitive verb, see, is twice marked, being marked again by the minor morpheme gī. The goal or object of this transitive verb is given in the topic in which he is not pivotal (not I saw him as he was writing) but exclusively actor; it is the topic as a whole that functions as the object or goal of the sentence, rather than the person of the topic, he, which is followed by a minor morpheme specifying

that the person embedded in the topic is actor of a transitive verb, write, which precedes another minor morpheme marking goal for the topic as a whole.

In each house there are two rooms.

(9) k'aŋ pa rere la k'aŋ mi ŋyi ŋyi
[house masc. each inside] [house fem. (two two)
yö wa re
are pf. non-1st-person]

Note that house plus masculine formative means house, while house plus feminine formative means room. The topic is specified in the first phrase (as for each house inside), and the comment in the second (two rooms are distributed). The reduplicated modifier in parentheses after the modified, room, marks distributive plural.

Compare sentences (10) and (11):

I stayed.

(10) ŋ-a de wa yį̄
[I-m] [stay pf. 1st person non-present]

I stayed on and on.

(11) ŋ-a de de wa yį̄
[I-m] [(stay stay) pf. 1st person non-present]

Reduplicated verb nucleus marks continuative aspect.

Compare sentences (12) and (13).

I may go in future.

(12) ŋ-a ɛ'į̄ ɛ'o gi yį̄
[I-m] [go may imperfective 1st person non-present]

In the second, or comment, phrase, the verb is followed by three verb complements: modal, tense-aspect without person, and tense with person marked redun-

dantly. Person is also marked in the first or topic phrase.

I am about to go.

(13) ŋ̀-a ʧ'i̱ ʧ'o ʧ'o yi̱

[I-m] [go (may may) 1st person non-present]

In the second or comment phrase the verb is again followed by three verb complements, but the first of these (the modal complement) is reduplicated. This reduplication marks inceptive aspect (ready to, on the verge of), and the impending act is classified by the last verb complement as non-present (hence either past or future). There is possible ambiguity here: in a narrative context, the same sentence might be understood as I was about to go.

His eye signalled me.

(14) k'-oŋ gi ʧ'ē gĭ ŋ-ā da naŋ ʧ'uŋ

[he-m of eye tr. actor] [me-m] [gesture hon. pf.]

The morpheme for third person is followed by a suffix usually associated with marking actor of intransitive verb. In the first phrase here, however, he-m is followed by the morpheme glossed of, for relating person as possessor to following noun, glossed eye, before morpheme specifying that eye is the actor of transitive verb. The second phrase specifies the object of the sentence by morpheme for 1st person followed by suffix which has the same shape whether marking subject of intransitive verb or goal of transitive verb. The transitive verb for this sentence appears as the phrase nucleus of the final phrase, as modified before honorific modifier, and also before verb complement for perfective. The order of the phrases, S-O-V (subject-object-verb) serves to confirm the reading of me for 1st person marker plus suffix in the second phrase.

They will probably be able to come.

(15) k'-oŋ c'o yoŋ t'u dōk'apo re

[he-m pl.] [come able ⟨probably⟩ m]

The translation they is obtained from the glosses of the first phrase, morpheme for third person followed by suffix specifying that this person is actor of intransitive verb, and followed by pluralizer. The next phrase begins with phrase nucleus for come, and two verbal complements discontinuous because interrupted by the optional morpheme complex for the modal probably. The last minor morpheme ($\frac{m}{M}$) in this phrase marks person (non-1st) redundantly, and also marks tense (non-present, which could be past or future; the preceding modal complex for ⟨probably⟩ confirms the reading of future rather than past).

He watered the flowers.

(16) k'-ō metō la č'u lu soŋ

[he-m] [flower goal] [(water pour) pf.]

These three phrases follow the favorite order, S-O-V. The Subject is he before suffix (-m) which specifies that the actor is actor of a transitive verb. The Object is flower, explicitly marked as object by a goal marker in the same phrase, which also occurs in the favorite Object order (S-O-V). The Verb phrase includes a compound verb followed by a verb complement.

I hit him with my hand.

(17) ŋ-ē k'-ō lak p-ē duŋ ŋa yī

[I-m] [he-goal] [hand masc.-with] [hit pf. 1st-non-pres.]

These four phrases show an Instrumental (I) phrase before the invariably final V phrase is an order much like the favorite order but extended to S-O-I-V. The Subject is marked by the morpheme for I before suffix (-m), indicating that the actor is actor of a transitive verb (but that actor is instrumental actor in other sentences, as (18), below). The Object is marked by the morpheme for

3rd person, explicitly marked as goal by the following suffix. The Instrumental phrase includes the morpheme for hand, before masculine noun formative morpheme with suffix for instrumental, glossed with. The Verb phrase includes the morpheme for hit, followed by two verb complements, that for perfective, and that for 1st person non-present, which redundantly marks past after the verb complement for perfective.

Compare sentences (18) and (19) which follow.

This work can be done by me.

(18) leka di ŋ-ē č'e t'u pa du

[work this][me-by] [do able pf. non-1st-present]

These three phrases follow the order S-I-V. The Subject phrase exemplifies once again the modified (work)-modifier (this) order. The Instrumental phrase is marked by morpheme for 1st person before suffix for instrumental which is glossed -by. The Verb phrase includes as phrase nucleus the verb for do, followed by three verb complements, that for able, that for perfective and that for non-1st person-present. Hence, an alternative free translation, stressing the present, might be The work is done by me.

This work cannot be done by me.

(19) leka di ŋ-ē č'e t'u pa min-du

[work this] [me-by] [do able pf. not-m]

Here again, the three phrases follow the order, S-I-V. The difference between sentences (18) and (19) lies in the last 'word' of the third phrase, which marks not before suffix (-m) usually for non-1st person-present, but since only du (and not yö) occurs immediately after negative, the distinction between 1st and non-1st person is neutralized here. The free translation, therefore, might be that given above (by analogy to sentence (18)) or I can't do this work.

As for the owner of this white horse, who might he be?

(20) ta kar po tē-i dak po su re

[horse white m this-of own m] [who $\frac{M}{m}$]

The topic includes a modified-modifier sequence horse before the modifiers white and this—or rather this-of, which relates, in a possessive or genitive way, the modified horse to the owner. The same minor morpheme (m) occurs twice in this phrase, after the morpheme for white, where it serves as an adjective formative, and after the morpheme for own, where it serves as an adjective or noun formative. The comment phrase begins with an interrogative embedded subject, the morpheme for who, and is followed by a dual function morpheme which has the value of non-1st person non-present ($\frac{m}{M}$), when it occurs after a verb in other phrases. But here, where there is no other verb in the phrase, the dual function morpheme marks verb to be, to have ($\frac{M}{m}$) in addition to marking person and tense, as above.

These three good dogs belong to me.

(21) ŋ-a la k'i yak po sum te c'o yö

[1st person-m to] [dog good m three this pl.] [$\frac{M}{m}$]

In the first phrase, the morpheme for 1st person is followed by a suffix which marks either actor of intransitive verb or goal—the latter, in this phrase which includes the morpheme for goal or to goal. In the second phrase, the modified dog is followed by morphemes for good, adjective formative (m), three, and plural marker which extends from this (these) to dog (dogs). The dual function morpheme in the last phrase functions as a major morpheme here ($\frac{M}{m}$): belongs to 1st person present, but in other phrases with prior verb, the same morpheme functions as a verb complement ($\frac{m}{M}$) in which the 1st person reference is a redundant echo of the already specified actor. The phrase order of this sentence is

O-S-V, instead of the favorite S-O-V order.

The evil men are more numerous than the good ones.

(22) mi saŋ po lɛ̄ ŋɛ̄ pa maŋ gi re

[man good m than] [evil m] [many impf. $\frac{m}{M}$]

The first phrase shows modified man, before modifiers good and adjective formative (m), followed by the morpheme for than, which compares the first phrase with the second. In the second or subject phrase the nucleus evil is followed by agentive or masculine noun formative. In the final phrase, two verb complements follow the morpheme for many, that for imperfective, and that for non-1st non-present, which together lend an emphasis to the continuousness of the state—the combination is translatable as future in other phrases. The comparative phrase (C) is the initial one in a C-S-V order.

Red is better than white.

(23) kar po lɛ̄ mar po ǰ'e gi re

[white m than] [red m] [good impf. $\frac{m}{M}$]

This sentence, like (22) above, shows a C-S-V order (comparative phrase-subject-verb). The morpheme for good is followed by the same two verb complements, here again suggesting the continuousness of the fact.

Compare sentences (24) and (25) which follow. The final verb complement ($\frac{m}{M}$) of the preceding two sentences functions as a major morpheme ($\frac{M}{m}$) in the following two sentences—will be, in reference to non-1st person non-present.

This will be best.

(24) di yak sʊ̄ re

[this] [good most] [will be]

The second phrase, marking the superlative, occupies the same position as the Object phrase in the favorite S-O-V order.

This will be best of all.

(25) t'aŋ mē-i naŋ ne di yak s̈ö re

[all m-m middle from] [this] [good most] [will be]

The first phrase is a pre-topic phrase in which the phrase nucleus for all is followed by a feminine noun formative and genitive marker, and by other minor morphemes for place and position.

Compare position of the negative marker in the verb phrases of sentences (26) and (27): before transitive verb but after intransitive verb. This might be better stated in terms of what follows the negative marker, glossed not: either the phrase nucleus follows, or the verb complement follows.

He (or you) did not give a letter.

(26) č'arī ma naŋ ŋa re

[letter] [not give pf. $\frac{m}{M}$]

The phrase order is O-V (object-verb), even though the actor is specified in the last of the two verb complements (m) of the verb phrase: non-1st person, non-present.

He is not going.

(27) k'-o ḍo gi ma re

[he-m] [go imperfective not $\frac{m}{M}$]

The phrase order here is S-V (subject-verb). It is specified in the subject phrase that the actor is explicitly 3rd person and, by following suffix (-m), that the actor is actor of an intransitive verb. Person distinction is neutralized after negative in the final verb complement (m), as in sentence (19).

Compare sentences (28) and (29) which follow. Both show dependent clauses or phrases immediately after the topic phrase which is either subject

of both phrases or clauses in the comment, as in sentence (28), or object of both phrases or clauses in the comment, as in sentence (29). But, as for all sentences, the main verb in both sentences (28) and (29) is in the sentence final phrase.

He, after having mounted the horse, rode off.

(28) k'-ō ta ɣsō nē ʧ'im pa re
[he-m] [horse mount after] [ride off pf. $\bar{M}^{m}$]

As phrased here, the order is S-OV-V. But the medial OV (object-verb) can be regarded as a clause consisting of two phrases with [horse] as object phrase and [mounted after] as non-final verb phrase followed by sentence final verb phrase. Topic functions as subject in the first phrase in which the morpheme for 3rd person precedes a suffix (-m) which indicates that the actor is actor of a transitive verb, and the verb of the dependent clause which immediately follows the topic, as the first clause of the comment, does indeed include a transitive verb, mount, after the object, horse. In the main clause or sentence final phrase, the verb is intransitive, ride off, and as usual precedes verb complements, the second of which, $(\frac{m}{M})$, redundantly specifies that the actor is non-first person and that the tense is non-present.

As for this, he did it because I was not able to.

(29) di ŋ-ē ɣ'e ma t'u cą k'-oŋ gĭ naŋ ŋa re
[this] [I-m do not able because][he-m m do pf. $\frac{m}{M}$]

As phrased here, the order is O-SV-SV, with topic functioning as object of both the following clauses or phrases, each of which includes an embedded subject, but of different person. The subject marker in the medial phrase for I is followed by a suffix (-m), and the subject marker in the sentence final phrase for he is followed by -m as well as m; these minor morphemes

specify that the actors are actors of transitive verbs which are both glossed do, even though they are formally different, /č'e/ in the medial or dependent clause, and /naŋ/ in the final phrase or main verb clause. Classical Tibetan verb bases appear in as many as four forms, distinguishing present, past, future and imperative; modern colloquial Tibetan usually substitutes the past form of the base for all the other possible forms. The object of both verbs in this sentence is given in the topic, glossed this.

[There are certain inconsistencies in the sentences cited above, and the comments on them, which we have no way of resolving since they represent inconsistencies in the sources (e.g. k'-oŋ is cited as both 3rd person actor of transitive and as 3rd person actor of intransitive; both ŋ-ā and ŋ-a are cited as actor of intransitive or goal and the latter is also cited as actor of transitive).]

4.2. The language or dialect which Grierson and Konow call the sole member of an Eastern Tibetan group is called Khams, and is listed as in the Central group by Shafer, but is counted as a dialect of the West Tibetan language by us, because it is said to have the same characteristics as other dialects which are geographically western. The latter are also grouped together by Shafer, but under another name than our West Tibetan language; he calls them the West Bodish Unit of the Bodish Branch of the Bodish Section of the Bodic Division. Lest this detailed sub-classification, which uses the term 'Bodish' repeatedly, be confused with the language name 'Bodo', which occurs in another family of languages though in the same phylum with the Tibetan language family, it should be explained that Shafer obtains the technical term 'Bodish' by Anglicizing the self-designation term for Tibetan, namely Bhōṭia. In short, 'Tibetan' is derived from English while 'Bodish' is derived from a self-designating term meaning Tibetan. According to Shafer, all the dialects of his West Bodish are spoken in Kashmir. Additional information about dialects of the West Tibetan language now follows.

Baltī is indeed spoken in the Kashmir state of India, but the 1901 census of Kashmir used Baltī for languages spoken in Purik as well as in Baltistan, arriving at a total of 134,372 speakers. (We discuss Purik separately as another dialect.) The Baltī alone have recently been estimated to number 75,000. From a corpus of 636 syllables, it appears that Baltī consonant finals are /t k/ and voiced / g / among stops, / s ɣ/ among fricatives, and /n ŋ / among nasals. All these finals also occur as initials, in an expanded consonant system which is much the same for Baltī and other West Tibetan dialects (see below). The consonant initial clusters noted for

Baltī, specifically, are of the C_1 C_2 type in which C_2 is /r/ after some stops; C_2 is also a stop consonant after /s/ or /d/ . A first member of the cluster, C_1, is /s/ or /d/ before stop, or /l/ before some stops or /z/; and C_1 is also /r/ before stops, /s/, or /m/.

Ladakhī is spoken in the province of Ladakh, India, a part of the Ladakh Wazarat of Kashmir. The dialect area is called 'Great Poliu' or 'Great Tibet'. This is the dialect known as Bod Butun by Kašmirīs at the turn of the century. It is known as Mar-yul by Tibetans. The self-designations for this dialect are Budhī, as well as Ladakhī; and these names appear to be cover terms for additional sub-dialects known as

Sham, the dialect of Hanu, spoken in an area from the west to a line midway between Saspola and Basgo in the east;

Leh, a dialect spoken east of the Sham dialect, in an area stretching eastward almost as far as Sheh;

Rang, a dialect spoken east of the Leh dialect.

The differences among main dialects of the West Tibetan language include so many differences in words or morphemes that mutual intelligibility may turn out to be a consequence of learning a closely related language rather than adapting to a divergent dialect. But the difference between the sub-dialects of Ladakhī, as enumerated above, seems quite restricted to a difference in sounds and word construction, with much the same words occurring in all the subdialects.

Purik is spoken in the province of Purik, a part of Baltistan, from Mulbe to Dras. According to Bailey (1920), it is, as Grierson and Konow also say, closely related to Ladakhī and Baltī; it is classified as "a dialect of Tibetan belonging to the Tibeto-Himalayan branch of Tibeto-Burman languages" and is

more specifically located as "north and south, from the headwaters of the Surū and Dras Rivers down to a little below the junction of the Surū with the Indus at Kharman; west to east, it extends from Dras past Kargil to the pass near Maulba Chamba on the way to Teh." The distinctions among consonants listed phonetically include retroflex consonants for stops, fricative, and vibrant /ṛ/, and it is quite possible that the almost pan-Indian retroflex contrast extended distinctively to dialects of the West Tibetan language in an areal linguistic perspective.

Lahul is spoken along the headwaters of the Chandra and Bhaga Rivers almost to their junction--to within fifteen miles of their junction; it is also spoken beyond the Mid-Himalayas, in the mountainous parts of Pangī of Chanbā, according to an old source, Jaesche (1807). When described in modern terms, Lahul may well turn out to be as divergent a dialect of the West Tibetan language as it appears to be from such observations as that it is a 'link dialect' (a turn of the century anticipation of modern 'chain relationships' among dialects) between the West Tibetan language and the Central Tibetan language.

Kahm, Kahms, or Khams-yal, is spoken between the province of Ü and China, and is counted as a dialect of the West Tibetan language (despite its relatively eastern geographic location) because it shares with the West Tibetan language features which are lacking in the Central Tibetan language.

There is relatively little phonetic information on the dialects of the West Tibetan language; and in the earlier work it is not always indicated where sounds are distinctive or contrastive. By combining more or less fragmentary information from the West Tibetan dialects, it seems likely that the following distinctions were made in most:

p	t	c	ṭ	č	k	ʔ
p'	t'	c'	ṭ'	č'	k'	
b	d	ʒ	ḍ	ǯ	g	
		s	ṣ	š		h
		z		ž		ɣ
m	n			n^y	ŋ	
	l ɬ					
	r		ṛ (or r^h)			
w	y					

The vowel systems of Baltī, Purik, and Ladakhī all seem to be of the 2 (FB) over N type because front-back contrasts are made at two tongue heights, high /i u/ and mid / e o/; but there is a neutral vowel—i.e. with no front-back contrast—at low tongue-height, /a/. Length contrast distinguishes short from long for all five vowels; and there is no tone contrast.

In the same West Tibetan language there is another dialect, Lahul, which appears to share the same 2 (FB) over N vowel type, with long-short contrast and lack of tone contrast. In speaking of mutual intelligibility among dialects of the West Tibetan language, Baltī, Purik and Ladakhī speakers are often specifically mentioned as understanding each other with ease. The lack of explicit inclusion of Lahul suggests that if this dialect is intelligible with the other three, it is so with difficulty.

However, in a few divergences from the common consonant distinctions as charted above for the West Tibetan language, Baltī as well as Lahul differs in some respects. Baltī, for example, distinguishes /v/ from the three fricatives as charted, making four distinctions for fricatives. Both Baltī and Lahūl

make three additional distinctions among stop consonants, by virtue of contrasting /k/, as charted, with palatalized /k^y/, /g/, as charted, with /g^y/, and aspirated /k^h/, as charted, with palatalized-aspirated /k^{yh}/. On the other hand, Baltī, like Purik and Ladakhī, distinguishes two r sounds (/r/ from retroflex /ṛ/ in some dialects, and /r/ from aspirated /r^h/ in other dialects). However, there is no contrast of r sounds in Lahul—only one /r/; nor is there a contrast between laterals in Lahul—only one /l/; and there are only two (rather than three) contrasts between Lahul voiceless sibilants, /s š/.

4.3. Among the dialects of the Central Tibetan language, those which are clearly localized are generally categorized as non-Lhasa Tibetan, while the Lhasa dialect is less clearly localized, perhaps because it is in a sense the lingua franca of Tibet, and also because it is, in another sense, a vertical or class dialect rather than a horizontal or regional dialect. And there may well be sub-dialects of the Lhasa dialect, which has its geographic center in the city of Lhasa where it is the language of officialdom; but of course officialdom extends beyond the city of Lhasa, and even beyond the Ü province in which the city of Lhasa is located. Even within the Ü province, non-Lhasa dialects of the Central Tibetan language are spoken.

Degrees of politeness in spoken Lhasa distinguish three styles:
an ordinary style, used between people of lower social rank ('inferiors' speaking to 'inferiors');
an honorific style, used between people of higher social rank (and also used by 'inferiors' in speaking to a 'superior' or about a person of good social position);
the high-honorific style is reserved for speaking about or addressing directly persons in the very highest echelon of the Tibetan heirarchy, as the Dalai

Lama or a Pancheon Lama.

Lhasa has a written literature which extends back to the 7th and 8th century AD; there is said to be some 'spelling pronunciation' among speakers who read this literature. The literate today use a kind of writing called yig skad, /yi-ke/, in correspondence.

In short, there is a reading style and a spelling style in addition to the three degrees of politeness distinguished in the speaking styles. Sprigg (1954) points out that each style requires a separate statement for the sounds which it distinguishes, for the grammatical rules which it follows, and for its lexical resources. Those who speak the Lhasa dialect are members of the sku-drag class in Lhasa and elsewhere, or members of other classes who associate with the elite sku-drag, especially in their formative years. Thus, Sprigg's informant, Righdyin I. Basipo, though born in Lhasa, was not himself a sku-drag; however, since he went to school with members of the sku-drag class, he became a speaker of Lhasa Tibetan.

In speaking this dialect, styles in the different degrees of politeness are different both by virtue of making different selections among morphemes, and by virtue of adding morphemes. Thus, honorific noun forms are different than the synonymous noun forms in the ordinary style, and a classifier is used before the noun in the honorific style, but not in the ordinary style. This means, for example, that the nouns for <u>song</u> (šē) and for <u>shoulder</u> (puŋpa) in the ordinary style may be used without preceding classifier. In the honorific style, the classifier for <u>body part</u> is ku-, and the classifier for <u>vocal sound</u> (in talking or singing) is suŋ; then, also, the noun form after the classifier is slightly different for <u>song</u> (suŋ šē) and for <u>shoulder</u> (ku puŋ).

A suffix used in polite address (-la), is simply omitted in the ordinary style. Selections are possible among what are classified as verbal auxilaries or verb particles, functioning synonymously but distinguishing ordinary style from honorific and either of these styles from high-honorific. So also similar selection among synonymous adverbs distinguishes ordinary style from honorific style.

Within the Central Tibetan language, the Lhasa dialect no doubt has independent status. It distinguishes one series of stops, rather than more than one, as in the Western Tibetan language (4.2, above), and as in the other Central Tibetan dialects generally (4.4, below). Both Miller (1955), and K'un Chang and B. Shefts (1961) agree on the single series of stops in the Lhasa Tibetan dialect, and agree that distinctions made by these stops include /p t ṭ č k ʔ/; additional affricates beside /č/, and a fronted /k/ beside /k/ may possibly be restricted to certain styles within the Lhasa dialect. There is also partial agreement on fricatives /s š h/, beside some additional fricatives; and on nasals /m n ñ ŋ/, beside a nasalizing phoneme; and on oral voiced continuants /l r/, beside aspirated /ř/ and the semivowels /w y/. So also for the vowel system, there is agreement in high tongue height distinctions /i ü u/; mid-tongue height /e ö o/, beside /ə/; and low tongue height /ɛ a/, beside /ɔ/. Oral vowels contrast with nasal vowels, and high tone contrasts at least with low tone. This generalized characterization of sound distinctions in the Lhasa Tibetan dialect may be compared with the prosodic description by Sprigg (1955), which distinguishes four affricates, including one with lateral off-glide, and obtains two series of stops by regarding aspiration and non-aspiration each to be a 'marker exponent', while voicing is not. Voiced stops are not distinctive ("voice occurs as an exponent of either the inter or

intra syllabic junction").

4.4. The other dialects of the Central Tibetan languages may be called collectively the non-Lhasa dialects. In Shafer's usage of 'Bodish section', there are some 'branches', as his Tsangla, Rgyarong (Pati, Wassu), and Gurung (Murmi, Thaksaya) which include languages (and dialects), as Gurung spoken in Central Nepal, which are not included in our Tibetan family. The latter is called by Shafer the 'Bodish branch of the Bodish section' and includes the 'West Bodish Unit' and an 'East Bodish Unit' which we have discussed under the rubric of the West Tibetan language (4.2, above). Shafer also includes under his 'Bodish branch of the Bodish section' two other'units', Central (Lhasa, etc.) and Southern (Sikkimese, etc.) which we discuss as various dialects of the Central Tibetan language whether Lhasa dialect (4.3, above) or non-Lhasa dialects.

The difficulty of making distinctions in the Tibetan family is apparent in language census data which lumps speakers of languages and dialects, so long as they speak any Tibetan, and then does not specify what languages and dialects are lumped in the Tibetan family. The language census figures are apparently contradictory. Thus, for 1956, there are listed, in a book giving lists of minorities in China, 2,770,000 Tibetan speakers, counting both those who live in western Szechwan and those who live in Tibet. In another source, the total population for Tibet alone, before 1947, is given as 1,000,000. The highest estimate that we have is that provided by the Center for Applied Linguistics: 6,000,000 as the total for all Tibetan speakers. This may include not only speakers of the West Tibetan language and the Lhasa and non-Lhasa dialects of the Central Tibetan language, included here as members of the

Tibetan family, but also other languages which we discuss in a subsequent section as members of another family, Gyarung-Mishmi, though no doubt related to the Tibetan family. There may well be six million non-Burman (i.e. non-Burmese-Lolo, Bodo-Naga-Kachin, Naga-Chin, Karen) speakers in South Asia and Tibet who do speak Sino-Tibetan languages, other than the Sino-Tibetan languages in East Asia and Southeast Asia. The language census estimates, then, are not necessarily conflicting, but suffer from using the label 'Tibetan' as a cover term for one or more than one language family— possibly even for phylum assemblages of several language families. In a Japanese source (Introduction to Languages of the World), the high estimate is given: four to six million Tibetan speakers.

We simply do not know what proportion of these millions of speakers under the cover-term 'Tibetan' speak the numerous dialects of the Central Tibetan language in Central and Southern Tibet, in north India, in Nepal, in Bhutan, and in China. Besides Lhasa (4.3, above) the dialect or language names for Central Tibetan include those spoken in the Ü and Tsang provinces of Tibet, and tribal or place names or self-designation names, as Sharpa, Kāgate, Däjongkä (also called Bhoṭia of Sikkim), Lhoke (also called Bhōṭia of Bhutan), Spiti, Jad, Garhwal, Nyamkat (Mnyamskad). In addition, Shafer includes Khams, possibly because it is geographically closer to the Central than to other West Tibetan dialects (4.2, above), and a list of Central Tibetan names for which we have no linguistic information; Hanniu, Ãba (Batang), Choni, Tseku, Dartsemdo (Tatsienlu), Nganshuenkuan, Sotati-po, Paurong, Dru, Panakha, Panags, Nyarong, Ngamdo (Amdo). The census estimates for particular dialects, below, are from Grierson, (1909), and apply to the number of speakers at the turn of the century (1901), unless otherwise stated.

The Garhwal dialect is spoken in the Painkhonda of Garhwal, India, by a few thousand speakers (the 1901 survey gives 1,891; a 1909 estimate gives 4,300). This dialect is said to be linguistically similar to that of Jad̤ of Tehri Garwal. Consonant initials distinguish aspirated from unaspirated, but not from voiced, which are said to have merged with voiceless aspirates.

The Nyamkat dialect, spoken in 1909 by 1544 speakers, is sometimes called Jad̤ (see below). Nyamkat means 'the language of equals' in Classical Tibetan; alternate names are Bad-kat, Bod-skad; Sangyas (Sangs-rgyas), the dialect of Buddhists. It is said to be linguistically closer to Spiti and Jad̤ than to Kanāwarī. As in Spiti, four series of stops are distinguished in Nyamkat—that is, the voiced as well as the voiceless stops distinguish aspirated from unaspirated.

The Jad̤ dialect is spoken in Tehri Garhwal, India, by one or two hundred speakers (the 1901 census gives 204; the 1909 estimate, 106). Jad̤ is said to be similar to the Spiti dialect (see below); as in Spiti and Nyamkat, four series of stops are distinguished.

The Spiti dialect is said to bear some resemblance to one of the West Tibetan dialects (Ladakhī), but is classified as a dialect of the Central Tibetan language because it distinguishes morphemes by tone. The voiceless unaspirated stops, / p t c ṭ č k^y k/, are matched completely by voiceless aspirated stops, and almost completely by voiced aspirated stops (there are no voiced aspirated affricates to match /c č /), and almost completely by unaspirated voiced stops (there is no voiced unaspirated stop to match the retroflex /ṭ/). The four series of stops occur as consonant initials; though stops occur as finals, the final stops are all voiceless unaspirated.

The Sharpa dialect name is a geographic designation (Shar east, Shar-pa inhabitant of an eastern country) for the Bhōṭia of Northeastern Nepal. Sharpa is spoken in Darjeeling, Sikkim, and Jalpaigun. It is said that Sharpa resembles Dän̈jongkä̈ more than Kāgate(see below). The Sharpa stops distinguish voiced stops (but not voiced aspirated stops) from aspirated as well as unaspirated voiceless stops, as do also the Dän̈jongkä̈ and Kāgate dialects.

The speakers of Dän̈jongkä̈ are the Bhōṭia of Sikkim. The district they live in is called Sikkim and Dän̈jong; the latter is translated rice district, and to this geographic name, the dialect name simply adds kä̈ for language (Dän̈jongkä̈ rice district language). The 1909 estimate of the number of Dän̈jongkä̈ speakers is much greater (20,000) than the 1901 census (8,825).

The Kāgate dialect is spoken in Eastern Nepal and Darjeeling. Both Kāgate and Dän̈jongkä̈ (above), distinguish three series of stops--voiceless aspirated and unaspirated from voiced stops.

The speakers of Lhoke are the Bhōṭia of Bhutan. The dialect name, Lhoke, is derived from a geographic name, the state of Bhutan, or Lho south; an inhabitant is called Lhopa, and the language Lhoke. Another name for Bhutan is Ḍukpa; hence, the Lhoke dialect speakers are also known as the Bhōṭia of Ḍukpa. The 1901 census gives more Lhoke speakers (8,980) than the 1909 estimate (5,079). Four series of stops are distinguished in Lokhe, as in Nyamkat, Jaḍ, and Spiti (see above).

The Central Tibetan dialect proper--the dialect spoken in the provinces of Ü and Tsang in Tibet--is the non-Lhasa colloquial standard, the lingua franca of Tibet, but spoken also in India, Nepal, Sikkim, and Bhutan. It is safer to use the cumbersome term, Central Tibetan dialect proper, rather than terms like Bhōṭiō, Huniyā, Shalgno, Kazi, or Lama, which may possibly

be used for other Central Tibetan dialects than the one regarded as a standard. Grierson gives only 7,968 speakers of the Central Tibetan dialect proper. Three series of stops are distinguished in this dialect, as in Sharpa, Ḍ̈ajongkä̈, and Kȟgate (see above).

Miller (1955) gives a more detailed description of the sounds of the Central Tibetan dialect proper, "as spoken by non-Lhasas of central and especially south central Tibet." In addition to the glottal stop, /ʔ/, seven oral stops are distinguished, labial /p/, velar /k/, and including the affricates, five stops produced by the front of the tongue: tongue-tip /t ṭ c/ and tongue-blade /t^{y} č/. These seven unaspirated stops are matched by seven aspirated stops, also voiceless, and by seven voiced stops. The fricatives are /s š h/, the nasals are /m n n^{y} ŋ/, and the semivowels /w y/. The liquids distinguish voiced from voiceless or aspirated: lateral /l̥/, and vibrant /r ř/. The vowel system is of the 2 (FF^{o}B) over N type because it distinguishes front, front rounded, and back vowel at two tongue heights, high /i ü u/ and mid /e ö o/; and for /a/ at low tongue height there is no much contrast (neutralized). Two tones, high and low, distinguish morphemes, and provide the determining environment for the sound of consonant finals. Thus any velar stop after a high tone is heard as a voiceless stop with fricative off-glide, [k^{y}], and after low tone as voiced. So also any labial stop after high tone is voiceless, with fricative off-glide, [p$^{\phi}$], and after low tone is voiced. Consonant finals include these labial and velar stops, the two voiced liquids, the nasals, and /s/. The initials include all the consonants, but only consonants, since vowels are non-initial.

5.0. In a language phylum perspective, it may be said that the Tibetan family is coordinate with the Gyarung-Mishmi language family. Or, what amounts to much the same thing, instead of speaking of two language families that are remotely related, it is possible to speak of coordinate branches of one language family. Following the latter way of speaking, Grierson-Konow classifies our Gyarung-Mishmi family as two additional branches of Tibetan languages, the Himalayan Tibetan languages and the North Assam Tibetan languages. Shafer's classification is in terms of branches and sections; Shafer feels certain that a few of his 'sections' (2 to 4, below) fall under 'Bodic', but that most of his 'sections' (5 to 10, below) are possibly Burmic, though probably 'Bodic'. In our phylum perspective, we merely note that the Gyarung-Mishmi language family is in the Sino-Tibetan phylum without placing it in a branching or sectional relation to a particular part of this phylum, as Shafer does:

(1) Rgyarung (presumably our Gyarung) and Gurung branches of the Bodish Section of Bodic;

(2) West Himalayish Section of Bodic;

(3) West Central Himalayish Section of Bodic;

(4) East Himalayish Section of Bodic;

(5) Newarish Section;

(6) Digarish Section;

(7) Midźuish Section;

(8) Hrusish Section;

(9) Dhimalish Section;

(10) Miśingish Section.

All the names in the list, (1) to (10) above, are included in our Gyarung-

Mishmi family. In addition to the majority of names in that list, (5) to (10) above, which Shafer leaves in some doubt as to their being 'Bodic' since there is a possibility that they may instead be 'Burmic', Shafer adds a few additional names which we do not include in the Gyarung-Mishmi family:

(11) Dzorgai;

(12) Kortse;

(13) "outer Manste";

(14) Pingfang.

These last, (11) to (14) above, are said to be either 'Bodic' or else 'Burmic' or even possibly representative of a separate division of 'Dzorgiac' or 'Dzorgaish', spoken in the Northeastern Tibetan Plateau, and in Szechwan and in Kansu.

In the sections below (5.1 ff), some languages are listed which are said to be 'non-pronominalizing' in contrast to others which have the fixed order V-sf.-VC (verb-suffix marking actor-verb complements) in the verb phrase. It is supposed that those languages which are 'non-pronominalizing' were influenced to be so by Tibetan languages in which the favorite phrase order within the sentence is S-O-V (subject-object-verb), and within the verb phrase, VN-VC (verb nucleus-verb complements), with the final verb complement redundantly marking or echoing the 'pronoun' (S) in the first phrase of the sentences, as well as marking tense. Our sample of Tibetan sentences (4.1, above) gives many instances in which a 'pronoun' is included in the verb phrase—not as a suffix after the verb nucleus, but as the last verb complement in the phrase which may include other particles to mark negative, clause dependency, and so on.

Some languages listed below distinguish gender (animate from inanimate);

and number (dual from plural, and in either number, exclusive from inclusive; and all these numbers from singular); count in the vigesimal rather than in the decimal system; and, as already noted, suffix subject or actor marker to verb before verb complements. In general, these distinctions and specializations are not generally found in other language families in the Sino-Tibetan phylum; conversely, they are found in some neighboring language families, as Munda, which is not a member of the Sino-Tibetan phylum.

5.1. The so-called 'western complex' of our Gyarung-Mishmi family is largely but not entirely a grouping of languages in terms of the geographic location of their speakers; it is possible to add to this a summary of typological samenesses (and differences) among the languages which are here included in the 'western complex' of what Grierson-Konow labels Himalayan-Pronominalized-Western. These languages largely overlap Shafer's West Himalayish Section--languages which are for the most part scattered in western Himalaya; but Shafer also includes in this group two languages which Grierson places (perhaps only for geographic reasons) in his Himalayan-Pronominalized-Eastern group (5.2, below).

Some of the 'western complex' languages distinguish glottalized stop consonants, /pˀ tˀ cˀ kˀ/, from other stops, as /p t c k/, and have other specialization in grammar, as already indicated (5.0, above). Their location is generally north of Almora, in Konowar, in Kangra, in Lahul, and in Chamba—all in western Himalaya.

Bunān is a language spoken along the Bhaga river, northeast for some 15 miles from its confluence with the Chandra. In its northernmost range, the Bunān speakers are neighbors of Spiti speakers (and Spiti is included in our Tibetan family, 4, above). In its southernmost range, and in the west,

the Bunān speakers are neighbors of Pāhāri speakers; the southern dialects are said to have been influenced by (borrowed from ?) Pāhāri, and these dialects are known by the language name Gāhrī, on the lower Bhaga River. Shafer pairs this Bunān language with another language, Thebor, and gives five additional names for 'dialects' of Thebor (Sumtśu, Zangram, Sungam, Kanam, Lippa). The grammar of Bunān includes suppletion for marking gender (feminine vs. masculine); and an extensive suffix inventory for marking case (goal or object case, ergative, instrumental, ablative, genitive and many positional cases—to, near, in, behind, from, etc.—additional suffixes or post positions following the genitive syntactic case); number (singular distinguished from plural for nouns, for the ergative case formation, and distinguished from dual in tense suffixes); tense of verbs (with the same suffix marking number and tense); part of speech formatives (suffixes for transforming nouns into verbs, verbs into nouns, for numerals, and the like); reduplication of adjective stems, with adjectival function marked also by relative order (modifier-modified); honorifics appear in the set of pronouns which distinguish number (with tense; and in a separate set for dual and plural—either exclusive or inclusive—and singular). The vigesimal system is used for higher numbers. Sound distinctions for all languages in the 'western complex', including Bunān, are summarized below.

Kanāwarī is spoken in the Sutlej Valley, near where the Spiti stream enters the valley. There are many alternate names for Kanāwarī, but especially Tibas skad; the two self-designations by the Kanāwarī people for themselves are Kanōrūg skadd and a slight variant, Kanōreunu skadd. There are other names, apparently restricted to certain dialects of the Kanāwarī people. In lower Kanawar, they are known locally as Milchang, Milchanany (Minchhang,

Minchhanäng), and Malhesti. Shafer includes the Kanāwarī language in his northwest branch of his West Himalayish Section, and adds two names; Tśitkhuli (or Chitkhuli) and Tukpa. The grammar of Kanāwarī, as spoken in central Bushahr, includes an extensive suffix inventory for marking case (goal or object case, nominative in general, actor of transitive verb, instrumental, genitive or possessive, and many position cases); number (singular contrasted with either dual or plural); paradigms of pronouns (with 'ordinary' distinguished from the 'respectful' for three persons, 1st, 2nd, 3rd; for three numbers, singular, dual, plural; for three cases, nominative, instrumental and genitive); verb voice and tense and person specification (intransitive, with no marker, vs. transitive, marked by object, vs. reflexive, marked by suffixes; present marked by finite or substantive suffixes but past by tense suffixes; person marker suffixes to verb distinguish the three numbers and inclusive from exclusive, as well as 'ordinary' from 'respectful'); modifier-modified order; reduplication of verbs in dependent clauses.

Lower Kanauri or Lower Kănawi is spoken within a hundred miles of Simla, and is presumably a dialect of Kanāwarī (above). Additional features of the grammar as given by Bailey (1920) include voice (passive and causative, marked by prefix, but no transitive actor), some mode but elaborate tense distinctions (distinguishing successive days from four days back to today to on the sixth day—beside general past, present, and future); aspect-tense (distinguishing present durative from past durative, for example, by elaborate suffix sequences, and from aspects or modes of 'necessity' and 'advisability'. The numeral system is vigesimal (20 plus 1 to 19 in addition; low number times 20 for multiplication up to 100). Sound distinctions made in the Kanāwarī and Lower Kanauri dialects are summarized below.

Kanāshī is spoken by a small tribe in a glen of the Bios or Bias Valley, living in an area about the village of Malāna or Mālānī; in 1901 there were 980 speakers. Grierson and his source (Franshawe) list various cases (transitive actor case, marked by suffix with noun before transitive verb, goal or object case suffixed to noun before imperative verb, and also benefactive goal, ablative, 'with' case, instrumental, locative or terminative case, and various positional cases); person marker paradigms distinguish an independent set from a set suffixed to verbs, three persons and singular and plural only, but distinguishing goal, benefactive, and possessive.

Paṭnī, also known as Manchāṭi is spoken in Lahul ("Manchat or Patan is a portion of British Lahul adjacent to Chamba Lahul"). The speakers numbered 2995. Their speech is said to be one of many Lahul dialects, and to be linguistically similar to Kanāwarī (see above). A given suffix may mark both number and case (as singular goal vs. plural goal or object; reduplication of stem final as well as a suffix marks genitive); person marking paradigms also distinguish number (dual and plural with exclusive-inclusive for each, and singular), as well as case (goal from actor from instrumental, and with a separate set for possessives); tense distinctions include present, past, compound past, and future.

Another Lahul dialect, known as Chamba Lāhuḷī, appears to Bailey (1905) to be "almost identical with Monchāṭi" (see above). Gender is marked suppletively. The numeral system is vigesimal.

Another dialect that is also said to resemble Manchaṭī is known as Ranglōī and Gōndlā and Tinan. This dialect was spoken some 15 miles from the confluence of the Bhaga River and the Chandra River, on the banks of the latter, and is said to have been "superseded by the Tibetan dialect spoken in

Spiti" (see 4, above).

Something like a composite picture is now given of the descriptions of sounds in the 'western complex' languages and dialects. In general, the early 20th century descriptions of consonants come closer to what is later confirmed for consonant distinctions than for vowel distinctions.

Aside from the glottal stop /ʔ/, which is reported for Kanāwārī but not for Bunān, for example, two affricate stops were distinguished, /c č/, beside /p t ṭ k/. All six oral stops were generally matched with aspirated stops and voiced stops; when a hole in the matching pattern is found, it is that one of the affricates or the retroflex /ṭ/ is not found in the aspirated or voiced series of stops. The fricatives distinguish /s $š^h$/ and one or both of the voiceless fricatives are matched with voiced /z ž/. Four nasals are distinguished, /m n ŋ/ and either retroflex /ṇ/ or palatalized /ñ/. Two each of liquids, /r l/, and of semivowels, /w y/, are distinguished.

Rangkas or Saukiyā Khun speakers are distinguished by dialect and location. The Rangkas dialect is spoken to the west of Dārmiyā, another dialect or language. Saukiyā Khun is spoken in five villages--one in Malla Johar, and four in Malla Danpur. These, together with Dārmiyā, Chaudāngsī (Tśaudangsi) and Byāngsī, are grouped together by Shafer in an Almora Branch of his West Himalayish Section. The Dārmiyā is spoken in Almora ("The Patti of Darma forms part of the Pargana of Darma in Almora.") in an area facing Tibet to the north, facing the Chhipula peak and the mountain chain including the Panchachuli to the west; the south boundary is the line connecting the Chhipula peak and the Kali River; the east boundary is the mountain chain that includes Yirgnajung and inter-

venes between the Dārmiyā speakers and the Byanges Valley and Patti Chaudangs, where Chaudāngsī is spoken. The area of Chaudangs is twelve miles long and eight wide—altogether some 100 square miles, between Khela and Nirpaniyodhura, and between the Kali and Dhauli Rivers. Byāngsī is spoken in Paṭṭi Byangs in an area facing Tibet to the north, the Kali River to the east and the south, and a lateral mountain chain to the west ("culminating in Virgnajung and Paṭṭi Chaudangs).

All these neighboring languages and dialects (Rangkas or Saukiyā Khun, Dārmiyā, Chaudāngsī, and Byāngsī) have much in common in their grammars, which can be touched upon in passing; but the information available is not sufficiently detailed to say how they differ from other languages in the 'western complex' or how they differ from each other (in any salient way). Whenever 'adjectives' are mentioned they are said to be the prior member (in a modifier-modified order). Case is always mentioned, and listed for syntactic cases, paradigms distinguishing four to eight cases (ergative or instrumental, dative or goal of benefactive, genitive, ablative, locative or terminative, and sometimes also positional glosses are included here: on, before, near, towards). But more often morphemes marking near, with, together with, between, on, under, from, behind, between, inside, and so on, are classified as 'postpositions'; it is sometimes said (e.g. for Byāngsī) that some 'postpositions' are suffixed only to genitive nouns. Person marker paradigms often distinguish not only 1st, 2nd, and 3rd persons, and singular and plural for each person, but also three cases for each person and each number: nominative and genitive, plus either agenitive or ergative. Suffixes after verbs also mark both person and some tenses; but it is possible that some tenses may be marked by prefixes, as the Dārmiyā past and future. The numeral system, when spec-

ified, is said to be vigesimal.

Janggalī belongs to the 'western complex' geographically and linguistically. Speakers of Janggalī (jungle language) are known as wild Banmanush (wood-men) who live in the forested region of Chhipula, in Askot Malla. The Janggalī language is the only representative of Shafer's Dźangalli Branch of the West Himalayish Section.

Thāmi and Bhrāmu belong to the 'western complex' linguistically but not geographically, since they are neighbors of languages in the 'eastern complex' (5.2, below). For linguistic reasons, Shafer includes both these languages under his Eastern Branch of his West Himalayish Section, despite the location of the Brāmus among other tribes of the Nepal Tarai. Both Thāmi and Brāmu are no doubt 'pronominalized' Himalayan languages, distinguish some cases by suffixes, as actor of transitive verb, goal case, possessive or genitive, and also in, on, upon, with, and in Thāmi by 'post position' under, before, behind. The 'adjective', when mentioned, is said to be in the usual modifier-modified order. Person marker paradigms distinguish the three persons, the two numbers, and possessive from nominative (and also from agentive in Thāmi).

5.2. Other languages of the Gyarung-Mishmi family—but not all others (see 5.3, below)—that are not discussed under the 'western complex' (5.1, above)—are discussed here. Except for Thāmi and Bhrāmu, these languages constitute the group that Grierson-Konow would classify as (complex) 'pronominalized' languages of the eastern-subgroup, but within this major group, subgroups of languages are brought together (after Shafer).

The first subgroup of Eastern Nepal languages includes Limbu, Yākhā, Rai, and numerous Khambu dialects; also the non-pronominalized Sunwār or

Sunuwār, as well as some other languages or dialects known only by name. Four of these languages are classified by Grierson-Konow as 'pronominalized' languages of the eastern subgroup (Limbu, Yākhā, Rai, Khambu). Shafer's classification includes not only these but additional names under his East Himalayish Section: Bahing, Thulung, Tśaurasya, as well as a Dumi Unit (Dumi, Khaling) for the Western Branch; and beside a Khambu Unit (adding Kulung and Sangpang, the same language—and Natśhereng); also a Bontawa Unit (Rodong, Waling—Runtśhenbung, Kiranti, Dungmali; Lambitśhong which is the same as Tśhingtang though not spoken in the same locality; Lohorong which is the same as Balali spoken elsewhere).

Limbu as well as Yākhā (see below) are said to be divergent representatives of the East Himalaya languages. They are neighboring languages in Eastern Nepal. Limbu is spoken east of the Yākhās and southeast of the Khambus (see below), in an area flanked by two rivers in Nepal (the Oud-kosi and the Kauki). The term Limbu or Das Limbu refers to the ten sub-tribal divisions of Limbu social organization, but is not used by the Limbu themselves. The Tibetans use the term Monpa for the Limbu and other neighbors on the Nepal side of the Himalayas. The Limbu are called Tsong or Chang by the Lepchas of Nepal; and are called Subah or Suffah (for chief) by the Kiránti. The self-designation term used by the Limbu themselves is Yākthūngbā.

The consonant distinctions of the Limbu language are given for three dialects (Fēdopia, Fāgūrāī, and Tamarkhōlēā) which agree in distinguishing only four stops/ p t č k/ which are matched by aspirated stops and voiced stops generally (but incompletely and probably as non-distinctive variants of the voiced stops, by voiced aspirated /b^h/ at least). There is a single fricative /s/; or at most, an /s h/ contrast in the three dialects. All disting-

uish three nasals, /m n ŋ/; two liquids, /r l/; and two semivowels, /w y/. The grammar includes suppletive selection for masculine and feminine human gender as well as suffixes for distinguishing between sex gender of animals; suffixes also mark number (as plural), possessive, accompaniment, for sake of, before (in time and place), behind, and distinguish actor of intransitive verb from actor of transitive verb. Person marking paradigms distinguish three persons, and possessive from verb dependency person, and five numbers (singular, dual, plural inclusive, plural exclusive, and plural). Verbs distinguish a middle voice, presumably from an active and a passive voice, with selection among person marking paradigms for actor and goal, with some prefix-suffix concordance. There are elaborate rules of sound change for marking the negative, after suffixes ending in consonant and vowels before labial or non-labial initials. Verb complements mark aspect, mode, interrogative, and quotative.

Yākhā is spoken by a small tribe between the Singilela mountains and the Arun River—1366 speakers are given in the 1904 census, including those living in Darjeeling and Sikkim. The self-designation for the Yākhā is Yakthomba (which means yak-herds, and refers to their focus of interest prior to their arrival in Nepal), or more honorifically Rōi. The consonant distinctions of Yākhā may be greater than those of Limbu, by the addition of retroflex /ṭ ḍ/ to the stops, and possibly by the addition of a voiced aspirated series of stops. It is explicitly said that the higher numbers in Yākhā are counted in a vigesimal system. The person marking paradigm distinguishes nominative, genitive, and agentive—for 1st, 2nd, and 3rd persons, singular and plural. In general, not only are the sound distinctions in Limbu and Yākhā generally the same, but the grammars of the two languages are typologically close.

Rai speakers are located in a region between two rivers (the Dud Kosi and the Tanbor) in that part of Nepal which is inhabited by Jimdār and Yākhā tribes. There is some confusion between the name Jimdār and Khambu (see below); the latter were said in the 1901 census, to "often assume the title of Rai and claim to be the same or Jimdārs, but their pretentions are not admitted in Nepal." Rai is of course an honorific title. The Rai language distinguishes /p t č k/, with completely matching aspirated voiceless stops, voiced stops, aspirated voiced stops, and nasals—a most symmetrical system. The grammar of Rai appears to be much like that of Limbu and Yākhā, despite the generalization that the latter are divergent—perhaps divergent in sharing fewer eastern Himalayan cognates?

The Khambu language is remarkable for its many dialects. It seems likely that most of the 'additional names' which Shafer gives for his East Himalayish Section, are dialect names of the Khambu language. These are now listed in the following several paragraphs.

In 1909, the Khambu of the Darjeeling District, between the Sun Kosi and the Arun, were known by the tribal name of Kiranti, and the region they inhabited was known as Nōlōkh Kirānt, but this Kirānt country was also shared by the Limbus (see above). The Khambu lived on the southern spurs of the Himalayas, north and east of the Jimdārs and the Yākhās. Variations in dialects of Kambu have been classified, according to the geographic range of the speakers, into Hither, Middle, and Further Kirant areas.

Khambu of the Hither Kirant (Wallo Kirant) extend from Sun Kosi to Likhu, and comprise the Lōhōrang and Chhingtang septs.

Khambu of the Middle Kirant (M̃ājh Kirant) extend from the Likhu to the Arun rivers, and are associated with a dozen names, presumably place names

which may represent subdialects (Rūngchenbung, Rōdong, Dungmālī, Khāling, Dūmi, Sangpāng, Bālālī, Lāmbichhōng, Bāhing, Thūlung, Kūlung, Wālung, Nācherēng). Some of these names appear as major dialects or separate languages under Shafer's East Himalayish Section (see above).

Khambu of the Further Kirant (Pallo Kirant) extend from the Arun to the Michi and the Sungilela ridge, and are associated with the name of Chourāsya (or Chaurāshya) Khambus. Additional and related peoples and languages and dialects of the Bontāwā, have criss-crossing relationships, as will be noted below.

Khambu of the Darjeeling District are remarkably similar in consonant systems to their neighbors. The distinctions in stops for unaspirated voiceless /p t č k/ are completely matched by a voiced series, and voiced as well as the voiceless series are completely matched by aspirated series. Nasals distinguish /m n ŋ/, fricatives /s š h/, liquids /l r/, and semivowels /w y/. Noun suffixes mark cases, including nominative, genitive, actor of transitive verb, possessive, and plural, and so on, and postpositions or successive suffixes mark positional cases, as before, behind, under, as well as with, for. Adjectives appear before nouns in modifier-modified order. Numerals appear in sequence with generic classifier suffixes. Person marking paradigms distinguish three persons, two numbers, and two cases for each (nominative and genitive). Possessive pronouns may be prefixed. Suffixes after verbs mark both person and tense (present or past and future). Causative is marked by suffixes; negative by prefix as well as by suffix.

The Khāling dialect of Khambu is spoken in the Middle Kirant, in the hills between the Likhu and Arun Rivers. The person marking paradigm shows a distinction of dual beside singular and plural for all three persons, and

for 1st person dual a further distinction between inclusive and exclusive.

The Dūmi dialect of Khambu includes two person marking paradigms, one of which is for possessive relationships, and the other, like that of the Khāling dialect above, for other relations, as actor of verb. The dual with inclusive-exclusive for 1st person is distinguished from singular and plural in both paradigms.

The Chaurāshya or Chourasya dialect of Khambu spoken in Further Kirant (Pallo) is more divergent from the preceding Khambu dialects in lexical content than they are from each other. In sound distinctions and grammar it seems to be no more divergent.

The Lāmbichhōng dialect of Khambu is spoken in Middle Kirant by a sept of the Wāling Bontawas. Noun suffixes distinguish masculine from feminine; other suffixes after nouns function as gender classifiers for rational beings in contrast to all other noun referents. There are two person marking paradigms, that for possessive simply distinguishes the three persons, while the other paradigm distinguishes three numbers as well (singular, dual, plural), and also inclusive-exclusive for 1st person.

The Nāchherēng dialect of Khambu is spoken in the Middle Kirant. Here again, as above, there are two person marking paradigms. Suppletive selection of different nouns for the same referent, combined with different suffixes, distinguishes masculine from feminine.

The above give a fair indication of the variability in grammar among the Middle Kirant dialects of Khambu—very little indeed. Even briefer notes follow the names of the remaining dialects listed in Grierson.

Rōdōng or Chāmling dialect of Khambus: sex gender by suppletion plus suffix, one person marking paradigm used as substitutes for nouns, another

as substitutes for adjectives (possessives).

The Rūngchenbūng are a sub-tribe of the Bontāwa Khamous, whose dialect includes besides the many expected noun suffixes, a prefix marking plural; two noun sequences in possessor-possessed order; sex gender marked by suppletion and suffixes; also numeral classifier gender marked by suffixes, two for human beings, one for things; two person marking paradigms, the one for possessive relationship distinguishing singular from plural, the other distinguishing in addition dual for all persons, and dual inclusive versus dual exclusive for 1st person.

Wāling dialect of Khambu, marking sex gender, as above; and two person marking paradigms, as above.

Bālāli and Lōhōrong dialects of Khambu mark sex gender by suffixes without suppletion (wā-pā cock, wā-mā hen), and dual and plural nouns by suffixes (mīna-chi two men, mīna-mi plural men), beside dual (as well as inclusive and exclusive) and singular and plural in both the possessive and actor person marking paradigms.

Sāngpāng dialect of Khambu marks sex gender by suffixes, and genitive and dual noun by suffixes; both person marking paradigms distinguish dual from singular and plural; suffixes after imperative verbs distinguish dual from plural subject, and a suffix that may follow any verb marks 1st person singular.

The dialect spoken by the Kūlung clan of the Khambus marks sex gender both by suffixes and by suppletion; it includes two person marking paradigms, neither of which distinguish dual from singular and plural.

The Dungmālī dialect of Khambu distinguishes dual from singular and plural in both person marking paradigms; there is a 'human numeral classifier', but apparently no sex gender; dual for nouns is marked by suffixes, plural by

prefix or suffix, and dual and plural suffixes appear also after imperative verbs.

The Thūlung dialect of Khambu distinguishes sex gender by suffix; also by suffix, numeral classifier for animals and for humans; though both person marking paradigms distinguish only singular from plural without dual, there is a 1st person distinction for plural inclusive and plural exclusive.

The Chhingtāng dialect of Khambu is not a Middle Kirant dialect, as are the last several dialects listed above; it is a Hither Kirant dialect spoken between the Sunkhosi and Likhu. There appear again two person marking paradigms, but without dual or exclusive-inclusive distinctions: merely singular and plural for 1st, 2nd, and 3rd persons which substitute for possessive or other nouns. Sex gender is marked by suppletion and suffixes.

Sunwār or Sunuwār is of course not includable with the preceding languages of the eastern subgroup, if the criterion for inclusion is that of being 'pronominalized'; Sunwār belongs in Grierson's 'non-pronominalized' Himalayan group, according to his typological criterion, which means the possibility of having person marking suffixes after verbs as well as (redundantly) marking persons by selection from a noun substitute paradigm. There is a person marking paradigm in Sunwār which distinguishes 1st, 2nd and 3rd persons, and singular and plural of each, and two cases: nominative and genitive. Possessive is marked by a noun prefix. There are two favorite orders of Sunwār sentences, S-V-O (subject-verb-object) beside S-O-V. The consonant distinctions include /p t č k/ for stops with matching voiced stops, and both series also aspirated; in addition there is an aspirated retroflex /ṭ/. The fricatives are /v s h/ rather than /s š h/, but there are two liquids and two semivowels, as in other Himalayan languages in the eastern subgroup. Sūnwar is peculiar in its elaboration of medial consonant clusters--nasal-stop, liquid-

stop, stop-liquid, and stop or nasal-semivowel.

Unlike Sunwār (above), but like most other languages in the eastern subgroup, Vāyus is a 'pronominalized' language. So also is Kusūnda, which appears not to be included in Shafer, at least not by that name. Shafer classifies not only Vāyu and Chēpāng, but also Māgarī (a 'non-pronominalized' language) in his West Central Himalayish Section--because these languages are spoken in Central Nepal. Some sameness and differences among these languages are now noted.

Vāyu speakers (also known as Hāyus) are found in small villages on both banks of the Kosi river in the central region of the Himalayas from the great valley of Nepal to where the Kosi turns southward into the plains. The consonant distinctions of Vāyu are most like those of Sunwār, with four series of stops marking four distinctions in each series (but without retroflex aspirated dental, as in Sunwār); the Vāyu nasals match the stops at all points of distinction: /m n ɳ ŋ/. The Vāyu grammar includes sex gender marked by suffixes and suppletion, and three 'numeral classifiers' marked by suffixes for the classification of persons (with redundant gender) and irrational beings. Numerals above four are counted in hands, feet, and scores (vigesimal). The person marking paradigms (noun substitutes) are unusually elaborate, distinguishing nominative, genitive (with suppletive forms before suffixes), instrumental for singular, plural and dual (exclusive and inclusive) persons. Person markers suffixed to verbs are also elaborate (rather than being reduced to a few distinctions, as in most other 'pronominalized' languages), and mark at the same time distinctions of person and voice (causative), or person and tense, or person and case (actor vs. goal). Most affixes are suffixes, and the occasional prefixes include a negative marker--negative of imperative

verb in the case of Vāyu.

Chēpāng speakers inhabit the forests west of the Great Valley of Nepal. The two person marking paradigms distinguish possessive persons from other persons, but the two paradigms could be easily reduced to one set, with different suffixes.

Kusūnda speakers live in the same district as the Chēpāng (above). They distinguish sex gender by suffixes and by suppletion, and two paradigms for possessive and other persons, largely by suffixes. Singular, dual and plural first persons and non-first person are also marked by suffixes after active verbs.

Māgarī speakers were found in 1909 in most parts of Nepal, in Darjeeling, Sīkkim, Bengal, and in the Naga hills of Assam. Their easternmost village is Tannang, 41 miles west of Khatmandu, in Nepal. The speakers numbered 18,476 in 1909. Māgarī is typologically a 'non-pronominalized' language, but on the basis of cognates Shafer includes it in his West Central Himalayish Section. There are in Māgarī more than the average number of stop distinctions, according to a source cited by Grierson: /p t ṭ c č k/, matched by voiced stops, with aspiratèd stops matching both the voiceless and voiced stops. Five nasals are distinguished, /m n ṇ ñ ŋ/; and for fricatives, /z/ beside the usual voiceless /s š h/; and the usual two liquids and two semivowels. The vowel system can be typologized as 2 ($FF^{O}B$) over N because front-front rounded-back distinctions are made both at high tongue height, /i ü u/, and at mid tongue height, /i ö o/, but neutralized at low tongue height, /a/. Classifiers are marked by prefixes (for body parts, etc.).

Dhīmāl is typologically a 'pronominalized' language, but Shafer leaves its divisional classification uncertain ("probably sections of Bodic, possibly

of Burmic, certainly not of Baric"), merely grouping Toṭo with Dhimal as languages of the Dhimalish Section (uncertain Division), located in Darjeeling and Jalpaiguri. The grammar marks some distinctions often encountered in 'pronominalized' languages but shows only one paradigm of person markers (noun substitutes) for singular and plural of three persons; possessive is marked by a suffix, as is plural of noun, actor of transitive verb, goal, locative, and genitive (and when a noun is to be compared, a comparative suffix follows the genitive suffix). Two sets of demonstratives distinguish animate-inanimate gender. There are only four stop distinctions, /p t č k/, with the usual four series, not always fully matched. Four fricatives are distinguished, /v s š h/ (rather than the usual three), three nasals, /m n ŋ/, and as usual two liquids and two semivowels.

Tōṭō is spoken by 200 or fewer speakers in Jalpaiguri, and has been classified as a 'non-pronominalized' language. It differs in its consonant system from Dhīmāl. Aside from the glottal stop, /ʔ/, five stops are distinguished, /p t ṭ č k/, and these oral stops are matched by voiced stops in two rather than four series. There are the usual three affricates, /s š h/ rather than four. For nasals, liquids and semivowels, Tōṭō makes the same distinctions as Dhīmāl. In Tōṭō there is one person marking paradigm (noun substitutes) for nominative, possessive, and accusative cases of 1st, 2nd, and 3rd persons, singular and plural. Verb suffixes mark tense.

Thākya is a central Nepal language, geographically (as are Murmi and Gurung, grouped together with Thākya by Shafer as representing the Gurung Branch in his Bodic Division). Apparently not enough is known of Thākya grammar to determine whether or not it is a 'pronominalized' language. There are two person marking paradigms, for possessive persons and for other persons,

both distinguishing singular, dual, and plural by suffixes. Sex gender is marked by suppletion.

5.3. The remaining languages in the Gyarung-Mishmi family are 'non-pronominalized' Himalayan languages. Shafer does not utilize the typological distinction as to whether persons are marked (redundantly) after verbs as well as by independent person marking suffixes; hence the languages listed below are not brought together, except for occasional pairs which are placed under different Branches and even under different Sections (as above, where Murmi and Gurung are set up, with Thākya, in one Gurung Branch of the Bodish Section).

Gurung and Gyarung are not different names for the same language or set of dialects. The Gurung tribe is found all over Nepal, and has even spread to Darjeeling and Sikkim. The Gyarung dialect described by Wolfenden (1936) is spoken in Darjeeling. The geographic range of Murmi (also known as Tāmāng, Bhōtiā, Ishāng, and Sain) is not quite as extensive as that for Gurung, being confined to the Valley of Nepal, to Darjeeling, and Sikkim. It was estimated that there were 7,481 Gurung speakers outside of Nepal alone, and the 1901 census of India gives 32,167 speakers of Murmi. Though different languages, Gurung, Gyarung, and Murmi have an overlapping geographic range and all show more than the average distinctions among stops (at least a half dozen), which are matched by voiced stops, with aspirated series for voiceless stops and (except for Gyarung) for voiced stops also. Nasals also make more than average distinctions: /m n ñ ŋ/ in Gurung and Gyarung (plus a rounded velar nasal in Gyarung, which also distinguishes a rounded velar stop from unrounded velar stop), and aspirated /m^h n^h/ (as well as aspirated liquids) beside /m n ŋ/ in Murmi, which permits many consonant clusters, as

does Gurung. An extraordinary proliferation of prefixes is given for Gyarung—e.g. person possessive prefix before relative prefix before body part noun, while the relative prefix rarely appears before other nouns; tense of verbs is also marked by prefixes, as are the infinitive and the imperative and the causative; but tense is also marked by suffixes, as is the case of nouns (genitive, dative, ablative, locative, instrumental, toward) in Gyarung. Gurung and Murmi include few prefixes. The suffixes of Murmi and Gurung mark causative, actor or instrumental case, a goal case which also functions as a directional marker, a genitive and a locative or terminative case; and after the noun with case suffix there may appear one of a dozen 'postpositions' for locational cases and the like. In both Murmi and Gurung, and in a more limited way in Gyarung, there are person marking paradigms for 1st, 2nd, and 3rd persons, singular and plural; person-number is further distinguished by case (nominative, agentive, genitive), except in Gyarung which marks the relation of possessive person to noun possessed not by a genitive person marker, but by a possessive person prefix.

Newari is spoken in Central Nepal, in Paton, Bhatgoan, and most of the small towns, by thousands of speakers in 1909 (7,873) but by hundreds of thousands of speakers today, according to an estimate received from the Center for Applied Linguistics (400,000). This enormous discrepancy in language census is not well accounted for, even if the higher figure is taken to refer to Shafer's Newarish, which is a cover term for both Newari and Pahrī, since only 300 speakers of the latter are given for 1909. It is quite possible that about 8,000 people spoke Newari in Paton and Bhatgoan and smaller towns in 1909, but that there were many places where Newari was spoken that were not entered into the 1909 total; and the estimate of fifty times 8,000 for

Newari speakers includes all the towns in which Newari is spoken today, together with the possibility of quintupled population since 1909.

Newari has been a written language since the 14th century.

Pahrī (or Pahī or Paḍhī) is spoken in the hills of Central Nepal, and is spoken of as a 'sub-dialect of Newari' by Grierson, and together the two dialects may be called by one language name (after Shafer). This language is hard to classify, especially in a classificatory system which requires that a language be either Bodic or Burmic.

The consonant systems of Newari and Pahrī are much the same. Stops distinguish /p t č k/ in four series, as in many other Gyarung-Mishmi languages; so also fricatives distinguish /s š h/, and nasals distinguish /m n ŋ/ in both dialects, with an additional palatalized /ñ/ in the Newari dialect; an aspirated series of nasals partly matches the voiced nasals, and there are two liquids, and two semivowels. Both dialects show many consonant clusters, mainly stop-r or semivowel, r-stop, nasal-stop. In the grammar, gender (animate-inanimate) is distinguished by prefixes, as is the negative. But suffixes distinguish cases, number, numeral classifiers for things in general, trees and plants, weapons and implements, flowers, fruits, days. Person marking paradigms distinguish three persons, singular and plural, for nominative, actor of transitive verb, and possessive. Verbs are suffixed for tense, voice, and transformation from verb to noun.

Mishmi appears to be a tribal name for three colonies of speakers who number 15,000, according to a recent estimate made by the Wycliffe Bible Translators (Summer Institute of Linguistics). Grierson gives the general area occupied by the Mishmi as centering in the mountains north of the Assam Valley, bounded on the west by the Dibang River, and bounded on the east by

the Lama Valley at Dzayul (subprefecture of Lhasa). Mishmi colonies have been found as far south as the Nemlang River (off-shoot of the Irawaddy), and east of the Dopha Bhunn mountain, and up the Brahmaputra to the borders of Tibet. It is uncertain whether the so-called colonies—Chulikātā, Digāru, and Miju—represent dialects of Mishmi or separate Mishmi languages—probably separate languages. The Digāru constitute the basis of Shafer's Digarish Section of the Bodic Division, in northeast Assam and into Tibet, which includes also a language or dialect called Midu, not to be confused with Miju. The latter appears to be the basis of Shafer's Midźuish Section of Bodic, although the evidence for relationship in the case of both 'Sections' is said to be probable rather than conclusive.

Chulikātā Mĭshmi is a North Assam language. Fragments of information on the grammar include a vowel prefix which transforms nouns into adjectives, in the order modifier-modified; sex gender is marked, or at least feminine is distinguished from non-feminine. Mishmi is said to have phonemic tone.

Digāru Mĭshmi is another North Assam language, which marks sex gender and human-animal gender. In some sequences of two nouns, the first is in possessor relationship to the second; in others the second is the modifier of the first in a modified-modifier order. Suffixes or 'postpositions' mark case (in, with, from) of nouns, tense of verbs (past, future), as well as voice (imperative, 1st person plural exhortative, negative imperative, causative), and mode (disiderative, potential). A person marking paradigm distinguishes singular and plural of 1st, 2nd and 3rd persons.

Mījū is a third North Assam language that is less closely related to the other two than they are to each other. Beside a prefix adjective formative (and a prefix for negative of verbs), two nouns in sequence may function

in a modified-modifier order. There is also a prefix for marking kinship terms, and also a possible sequence of two nouns in which the first stands in a possessor relation to the second. The three person markers (noun substitutes) form their plural by the same suffix. Suffixes also mark tense of verbs (past, remote past, future), as well as mode (imperative and potential).

The remaining language names identify Hill Tribes along the unadministered Assam-Tibet border (according to information gathered by the Wycliffe Bible Translators). Grammatical information given is from Grierson.

Aka speakers live in the hills north of the Assam Valley, and are flanked by the Bhutan to the west, and by the Buruli river to the east, which forms a boundary between them and the Daflas of the Daflā hills. The Aka or Angka are so called by their neighbors; their self designation is Hrusso, and also Tenae. Shafer prefers the first of these self-designation terms, and uses it for deriving a Hrusish Section which may be Bodic ('probably'), it not Burmic ('possibly'). In Aka, sex gender distinguishes animate nouns, and hence criss-crosses with an animate-inanimate gender (the latter including nouns which are noncommittal as to sex gender); there is also animal classifier, marked by a prefix, and a prefix for negative verbs. The modifier-modified order may be included in a longer sequence of noun modifier-modified noun, in which the first noun functions as possessor of the second:

fu-grā	gro	dsimic
classifier-horse	white	saddle,

for 'The horse's white saddle ...'. There is an apparent modified-modifier order, but this may represent a topic comment sentence, (man good for 'The man is good') rather than a modified-modifier sequence within topic phrase or comment phrase. Suffixes mark case (goal or object) and positional case

(in, from, under, before, behind—and instrumental), and mark the comparative of verbs, and tense and aspect. A person marking paradigm (noun substitutes) distinguishes singular and plural of 1st, 2nd and 3rd persons. Aka is said to be the most divergent of the languages of the Hill Tribes on the Assam-Tibet border.

Miri is spoken east of the Daflā country, and is said to distinguish morphemes by tone. The Miri or Abor-Miri, who are now known by their self-designation, Adi, are supposed to be as closely related to Mishmi, where tone is also significant, as to the Daflās.

The Daflās live not only in the hills to the west of the Miri, but also in Darrang and Lakhimpur. In the Darrang district, the Daflās occupy the eastern part, in villages on the Bhareli River on the Darrang border, north of the Brahmaputra. The Daflās of Lakhimpur are called by others Bāngni, but their self-designation term is Nyī-sing, which means men of Sing. There is much dialect variation in the Dafla language. The grammar marks sex gender by suffix (and for animals by reduplication of the noun final syllable before the suffix; feminine, masculine). The numeral classifiers used are appropriate for domains of months, trees, animals, houses, villages, and eggs; they precede the quantifier. The noun case paradigm distinguishes nominative from suffixed forms for object, dative, locative, ablative, and vocative; positional cases are marked by subsequent suffixes or postpositions for inside, near, behind, etc. Tense is elaborate, and marked by suffixes after verbs for present, present definite, imperfective, future, past, past perfective. Mode, aspect, voice and negative are also marked by suffixes. In sequences of two nouns, the first functions as possessor of the second (e.g. elephant's leg), but in noun substitutes (person marking paradigm), the possessor is marked by suffix.

The Following Abbreviations Will Be Used

AA	American Anthropologist
ACLS . . .	American Council of Learned Societies
AES-P . . .	American Ethnological Society, Publication
AL	Anthropological Linguistics
APS-P . . .	American Philosophical Society, Proceedings
APS-T . . .	American Philosophical Society, Transactions
BAE-B . . .	Bureau of American Ethnology, Bulletin
BAE-R . . .	Bureau of American Ethnology, Report
CU	Columbia University Contributions to Anthropology
IJAL	International Journal of American Linguistics
IUPAL . . .	Indiana University Publications in Anthropology and Linguistics
JAF	Journal of American Folklore
JSAP	Journal de la Société des Américanistes de Paris
Lg	Language
RCPAFL . .	Research Center Publications in Anthropology, Folklore and Linguistics
SJA	Southwestern Journal of Anthropology
SIL	Studies in Linguistics
TCLP . . .	Travaux du Cercle Linguistique de Prague
UMPL . . .	University of Michigan Publications, Linguistics
UCPAAE . .	University of California Publications in American Archaeology and Ethnology
UCPL . . .	University of California Publications in Linguistics
VFPA . . .	Viking Fund Publications in Anthropology
WDWLS . .	William Dwight Whitney Linguistic Series

CONTENTS

SINO-TIBETAN FASCICLE FOUR

4.0. Scope of the Tibetan family 2

4.1. Sample of Tibetan sentences 6

4.2. West Tibetan 21

4.3. Lhasa Tibetan 25

4.4. Non-Lhasa Central Tibetan 28

5.0. Scope of the Gyarung-Mishmi family 33

5.1. Western complex languages 35

5.2. 'Pronominalized' languages 41

5.3. 'Non-pronominalized' languages 52